The PhyloCode: description and commentary

Peter L. Forey

*Department of Palaeontology, The Natural History Museum,
Cromwell Road, London SW7 5BD, U.K.*

Introduction

This essay is prompted by a seminar which I gave to the Department of
Palaeontology at The Natural History Museum, London. The size of the audience
indicated strong interest in the subject, and it was suggested that it would make a
suitable subject for an article in the *Bulletin of Zoological Nomenclature*. Since I was
critical of the proposals of the PhyloCode I was happy to agree to this on the
understanding that I could describe the aims and mechanics of the PhyloCode in as
neutral a way as possible while allowing myself the opportunity of personal
commentary. Therefore, this essay is divided into two distinct parts. Readers
may wish to cease reading at the end of the first part and form their own
opinions. The PhyloCode is published in preliminary form on the web at
www.ohio.edu/PhyloCode, and where possible I take direct quotes (designated in
italics — page numbers are irrelevant since different web download programs will
paginate differently) so as to avoid any personal filters beyond the selection from the
continuous text, which I encourage reading in total.

Part 1. The PhyloCode

The PhyloCode is a new system of Biological Nomenclature which is designed to
provide rules to govern the naming of clades across all of biology. The PhyloCode
is the formalisation of the ideas of Phylogenetic Nomenclature (also known as
phylogenetic taxonomy, see below) which has been discussed in a series of papers
beginning with De Queiroz & Gauthier (1990), although many of the issues raised by
advocates of Phylogenetic Nomenclature had been discussed long before. A near
comprehensive bibliography of Phylogenetic Nomenclature is given following the
Preface at the PhyloCode website. It has been discussed, refined and argued over in
three symposia, with the formal proposals being set out as a result of a meeting in
1998 at Harvard. The names of 26 people are attached to the PhyloCode as an
advisory group but it is unclear as to whether all of these are signatories to all of the
aims of the PhyloCode.

I should perhaps make it clear that terms such as Phylogenetic Taxonomy and
Phylogenetic Nomenclature were freely interchanged in the earlier papers on
Phylogenetic Nomenclature. The two are not the same. Phylogenetic Taxonomy is
effectively phylogenetic systematics. We can of course have Phylogenetic Taxonomy
without Phylogenetic Nomenclature.

Phylogenetic Nomenclature starts from the premise that there should be con-
gruence between phylogenetic hypotheses and nomenclature. At the moment it is
only in draft form, which may be perused at the web site cited above, and the authors
welcome comments as to its utility, practicality and the particulars. At present the
PhyloCode governs the naming of clades which may be previously un-named or

correspond to taxa above the species level in other biological Codes. Rules governing species names will be added in the future. '*The PhyloCode is designed so that it can be used concurrently with the preexisting Codes* [International Code of Zoological Nomenclature, International Code of Botanical Nomenclature and International Code of Bacteriological Nomenclature] *or (after rules governing species names are added) as the sole Code governing the names of taxa, if the scientific community ultimately decides that it should. The intent is not to replace existing names but to provide an alternative system for governing the application of both existing and newly proposed names.*' (Preface paragraph 3).

Fundamentally the PhyloCode is designed to name the various parts of the tree of life—clades (ultimately of species) — and it does this by explicit and sole reference to phylogeny. It runs counter to what we are all familiar with by giving no significance to ranks (Genus, Family, Order, etc.); it ignores familiar endings such as (in the Zoological Code) -idae for family, -inae for subfamily, -ini for tribe, etc. Such endings may be retained but they have no hierarchical significance, so that -ini may come to prescribe a more inclusive group than -idae.

The aims of the PhyloCode are directed toward reflecting phylogenetic hypotheses through a system of names and it emphasizes that the usage of those names should be explicit, unambiguous and stable: that is, they should not change their meaning through time. The PhyloCode defines names by reference to a hypothesised phylogeny but once a name is defined it may well be applicable in the context of other phylogenetic hypotheses.

The principles of the PhyloCode are stated under six headings (*PhyloCode Division I. Principles*):

'*1. Reference. The primary purpose of taxon names is to provide a means of referring to taxa, as opposed to indicating their characters, relationships, or membership.*

2. Clarity. Taxon names should be unambiguous in their designation of particular taxa. Nomenclatural clarity is achieved through explicit definitions.

3. Uniqueness. To promote clarity, each taxon should have only one accepted name, and each accepted name should refer to only one taxon.

4. Stability. The names of taxa should not change over time. As a corollary, it must be possible to name newly discovered taxa without changing the names of previously discovered taxa.

5. Phylogenetic context. The PhyloCode is concerned with the naming of taxa and the application of taxon names within a phylogenetic context.

6. The PhyloCode permits freedom of taxonomic opinion with regard to hypotheses about relationships; it only concerns how names are to be applied within the context of a given phylogenetic hypothesis.'

It needs to be pointed out here that 'taxon' refers to a clade or species. If a clade it does not matter how many species are included. Thus, a clade taxon may be what is referred to as a Genus or an Order or a Phylum under current Linnaean Taxonomy.

The PhyloCode recognises that there are three ways of naming a clade within a phylogenetic context and these lead to the explicit definitions referred to in the Principles. These are illustrated in Figure 1.

Consider a phylogeny as shown here in Figure 1a which shows a phylogeny leading to modern birds which are traditionally called Aves. This lineage may be considered
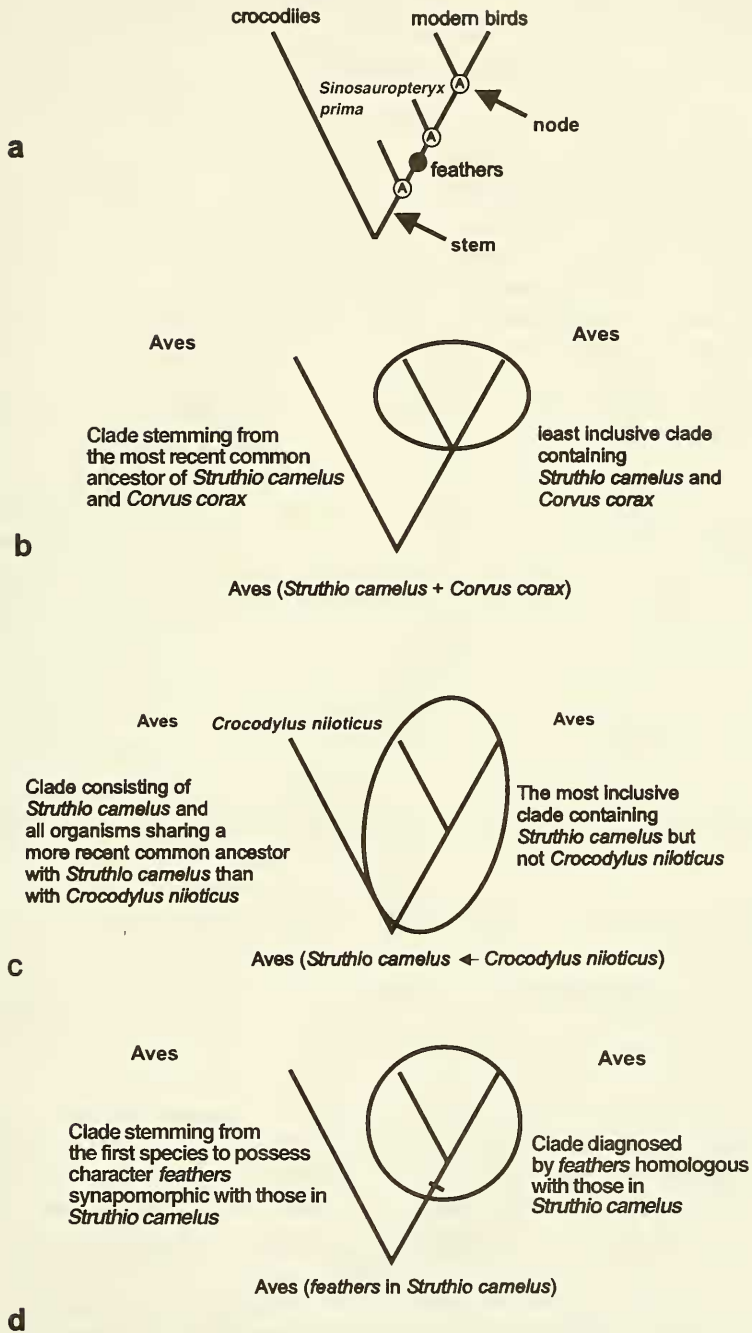
Figure 1. A phylogeny and the three definitions in which the PhyloCode suggests naming. The definitions to the left of each alternative include reference to ancestors, those on the right do not. Below each of the alternatives the shorthand code suggested by the PhyloCode is given. a. A phylogeny showing a lineage of taxa leading to birds and their modern sister-group. b. node based definition. c. Stem based definition. d. Apomorphy based definition.

as a series of cladogenic events, each split being marked as a node which was occupied by an ancestor 'A'. The intervening sections of the evolutionary history can be thought of as a series of stems. During the evolutionary history of the lineage changes between successive nodes may be characterised by the appearance of new characters (apomorphies) such as, in this case, feathers. This entire lineage will have a sister group, in this case designated as crocodiles. It needs to be pointed out to those readers more familiar with the crown, total and stem group concept of Hennig (see Jefferies, 1979) that there is partial overlap between phylogenetic systematics and Phylogenetic Nomenclature usage. In phylogenetic systematics there are the concepts of crown, total and stem groups. The crown group is the latest common ancestor plus all its descendants of a Recent group. The total group consists of all species more closely related to the crown group than to the Recent sister group and the stem group is the extinct paraphyletic assemblage leading up to the origin of the crown group.

In Phylogenetic Nomenclature there is no requirement that the node specify a crown group. In other words all crown groups are node-based groups but the converse is not true. Under Phylogenetic Nomenclature it is perfectly possible to recognise an entirely extinct node-based group. Similarly, all total groups are stem-based but not all stem-based clades are total groups. With this clarified we will continue within the terminology of Phylogenetic Nomenclature.

Under a node-based definition (Fig. 1b) the name 'Aves' is the name given to a clade stemming from the most recent common ancestor of (say) *Struthio camelus* and *Corvus corax*. Or, if we wish to strip out direct reference to ancestors, it may be expressed as on the right here as the least inclusive clade containing *Struthio camelus* and *Corvus corax*. The notation in parentheses below the tree is a suggestion for abbreviating the definition (Aves must have *Struthio camelus* **and** *Corvus corax*). *Struthio camelus* (the ostrich) and *Corvus corax* (the raven) are called specifiers. They serve exactly the same function as Linnaean types except their characters do not define the clade.

We could actually name as many birds to serve as specifiers as we wanted but two is the minimum. No matter what other birds such as sparrows, gannets or vultures are included, in this example the word Aves is constructed around the ostrich and the raven. Clade membership may expand or contract to include or exclude these extra taxa — and this depends upon the phylogeny — but the ostrich and raven must always be included. So in Phylogenetic Nomenclature we now have two types (specifiers).

In the stem-based definition (Fig. 1c) Aves is named as the clade consisting of *Struthio camelus* and all organisms sharing a more recent common ancestor with *Struthio camelus* than with *Crocodylus niloticus*. Or, again without specific reference to ancestors, as the most inclusive clade containing *Struthio camelus* but not *Crocodylus niloticus*. Here there is an included taxon—*Struthio camelus*—and an excluded taxon—*Crocodylus niloticus*. So again we have a reference to specifiers but this time one is included and one is specifically excluded. The stem-based definition states that a taxon is more closely related to one specifier or type than another. The shorthand notation is given beneath the diagram. So again we have two types (specifiers).

In the apomorphy-based definition (Fig. 1d) — with its abbreviation shown below the diagram — the definition is a clade stemming from the first species to possess the

character *feathers* synapomorphic with that in *Struthio camelus*. Or a clade diagnosed by *feathers* homologous with those in *Struthio camelus*. Here there is one specifier taxon and one specifier character. Two types (specifiers) but one is conceptually quite different from the other.

In order to name a clade there must be some phylogenetic hypothesis before us. Names are then applied in the context of that hypothesis. Should the hypothesis change then the taxonomic content implied by a name may change but the important point made by advocates of the PhyloCode is that the name is clear since it based on an explicit definition (stem-, node- or apomorphy-based), it is unique and stable since the taxon name is fixed to specifiers (taxa or characters).

To explain this, consider Figure 2 and the names Sarcopteryii and Choanata, and take the phylogeny of the left-hand column as the phylogeny current when the names Sarcopterygii and Choanata were coined. Under the PhyloCode the original author of the name would have had three choices of definition (node-, stem- or apomorphy-based) and choices of reference taxa. In this example let us say that the coelacanth and a frog (to represent a tetrapod) were used as specifiers for the node-based definition (Fig. $2A_1$) of Sarcopterygii and the lungfish and frog were used as specifiers for Choanata (the use of one anchor taxon — in this case the frog — for different definitions has been advocated by Lee, 1999a). Alternative phylogenies shown to the right (Figs. $2A_2$ and $2A_3$) would result in different taxon membership.

Let us say that the name had been introduced under the stem-based definition: that is, Sarcopterygii is the name given to the clade that includes the coelacanth but not the perch (an actinopterygian) and that Choanata is the name given to the clade including the frog but not the coelacanth (Fig. $2B_1$). The consequences of subsequent phylogenetic revisions are shown to the right (Figs. $2B_2$ and $2B_3$).

Lastly the same exercise (Figs. $2C_1$– $2C_3$) can be applied to apomorphy-based naming, with the exception that there would be ambiguity about the homology of fleshy fins under the second (Fig. $2C_2$) phylogenetic hypothesis so that the name Sarcopterygii could not be unambiguously applied in this case.

There are a number of features of this exercise to notice as properties of Phylogenetic Taxonomy. Firstly, a shift in taxon membership with changing ideas of phylogeny is perfectly acceptable to the PhyloCode since principle 1 states that: *The primary purpose of taxon names is to provide a means of referring to taxa, as opposed to indicating their characters, relationships, or* **membership** [my emphasis].

Second, ideas of relationships can vary substantially (e.g. the three theories given here) but, with one exception (Fig. $3C_2$, involving the apomorphy-based definition) there will always be some position at a node or along an internode on a phylogeny where the name Sarcopterygii will apply. That is also acceptable since principle 1 states: *The primary purpose of taxon names is to provide a means of referring to taxa, as opposed to indicating their characters,* **relationships**, *or membership.*

The third feature is that the name is applied to a phylogeny without reference to why that phylogeny should have been chosen. Again this is perfectly consistent with the aims of the PhyloCode: *The primary purpose of taxon names is to provide a means of referring to taxa, as opposed to indicating their* **characters**, *relationships, or membership.*

It is important to notice that a change in the phylogenetic hypothesis will cause a *different* change in the taxon membership and its hierarchical relationship to names
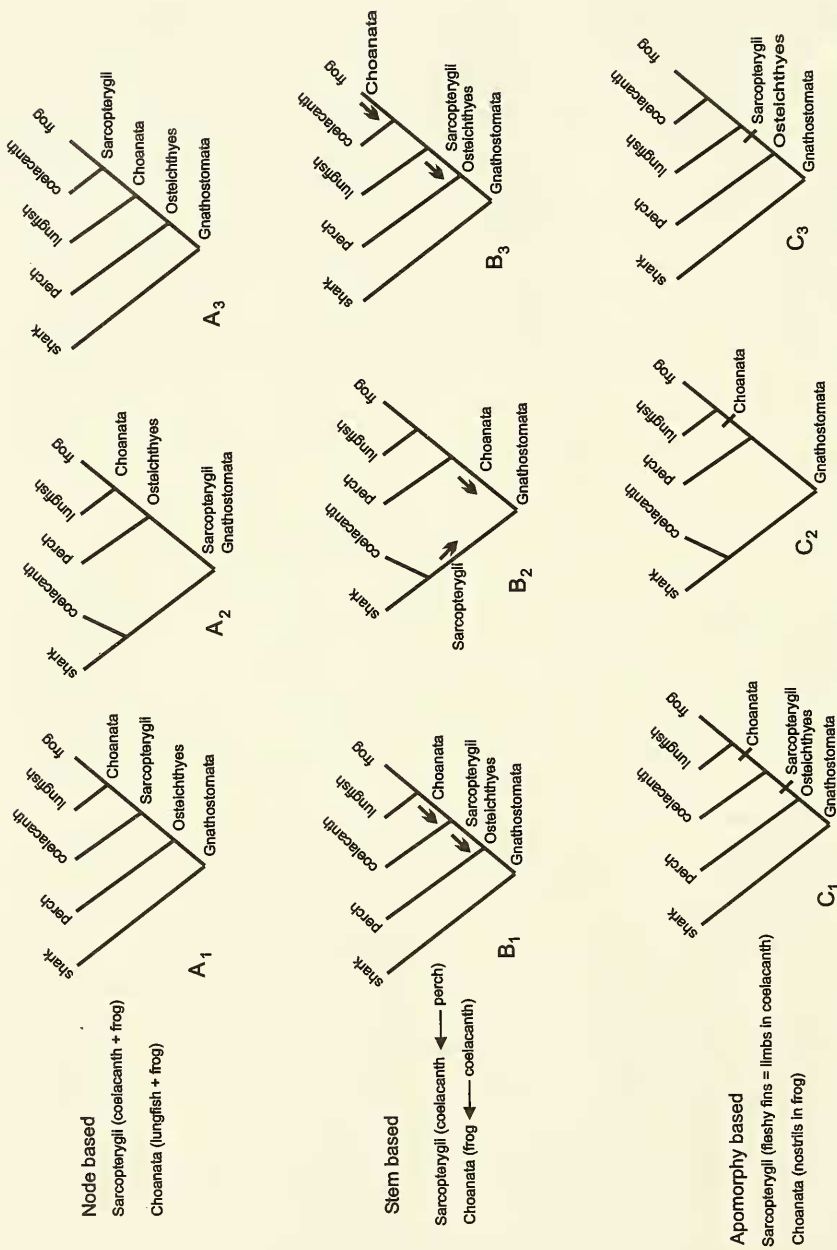
Figure 2. Stability and synonymy. In this diagram three hypotheses of the relationships of coelacanths to other vertebrates are given. That shown in the left column after Forey (1980), centre column after Lagios (1979), and that in the right column after Schultze (1987). The effects of the different taxonomic content of the names Choanata and Sarcopterygii can be mapped across each of the three definitions A — node based, B — stem based (arrows on trees are a convention used by Sereno, 1999), C — apomorphy based. Note that in C_2 and C_3 the names Sarcopterygii and Choanata could not be used because of problems of interpreting the specifier character as an homology (see text and Fig. 5 for discussion). In B_3 'Choanata' is co-extensive with the frog and presumably would not be named.
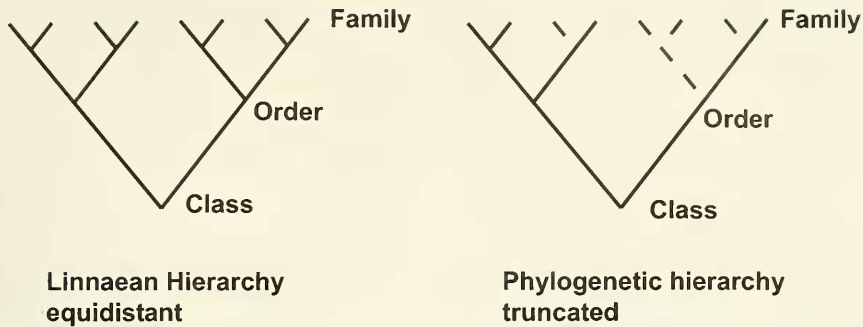
Figure 3. Truncated hierarchies and redundancy. A Linnaean hierarchy is symmetrical where all terminal taxa can be assigned inclusive rank. The hierarchy on the right approximates to many phylogenetic hypotheses where extinction (dotted lines) or reality means that some ranks for some taxa will be redundant.

of other clades depending on whether the name Sarcopterygii is node- or stem-based. Therefore, the PhyloCode makes it mandatory that the intended definition is stated when a name is proposed (see below).

Another phenomenon can be noted as a result of changing hypotheses when using the node-based name. This is the fact that the hierarchical relationships of names can reverse. Thus in Fig. $2A_1$ Choanata is more exclusive than Sarcopterygii whereas in Fig. $2A_3$ the reverse is true. Again, this is not a particular problem for the PhyloCode since it is not concerned with rank. However, the PhyloCode does suggest ways in which this and situations like it can be avoided: this is done by adding exclusion clauses or qualifiers to the definition. Thus, in this case we could say that the name Choanata is a name given to a clade including the lungfish and the frog but *excluding* the coelacanth. This would mean that in the phylogeny represented by Fig. $2A_3$ the name Choanata could not be used.

The final point to be outlined concerns synonymy and homonymy. To some extent these terms mean the same in the PhyloCode as in Linnaean Taxonomy. Thus homonymy is an instance where the same name is used for different taxa and synonymy is an instance where different names are used for the same taxon. But the meaning of homonymy has an additional dimension in the PhyloCode because of the different potential ways of defining a group (stem-, node- or apomorphy-based—see PhyloCode Note 13.2.3). With regard to synonymy there is the possibility of two names specifying the same taxon but since they may be defined in different ways (e.g. stem- and node-based) they may both be valid (PhyloCode, Note 14.1.2).

Practicalities and Governance

The PhyloCode will be part of the activities of '*The Society for Phylogenetic Nomenclature (SPN), an international, non-profit organization with no membership restrictions. Two committees of the SPN have responsibilities that pertain to this Code: the International Committee on Phylogenetic Nomenclature (ICPN) and the Registration Committee. [Note: These organizations do not yet exist. They will be established before the PhyloCode is implemented].*'

Thus the PhyloCode proposes a registration system whereby clade names are submitted electronically. In order to register a name certain pieces of information need to be provided (those marked with an asterisk being mandatory and others optional):

'Definition type (node based, stem based, apomorphy based, other . . .)*
*Phylogenetic definition**
List of specifiers, at least two being mandatory*
Qualifying clause
Reference phylogeny (bibliographic reference, URL, or Accession number in public repository)'

These then, are the aims and basic workings of the PhyloCode. Of course, there are many other provisions in the Code designed to streamline the naming process (orthography and authorship) and to deal with particular situations (e.g. hybrids). I encourage all to visit the PhyloCode website to read the full text.

Part 2. Commentary

Phylogenetic Nomenclature already has a history, with the main arguments and suggestions for its implementation having been put forward in a series of papers (e.g. De Queiroz & Gauthier, 1990, 1992, 1994; Rowe & Gauthier, 1992; Lee, 1998, 1999a,b; Sereno, 1999; Cantino, 2000). Counterviews have been expressed in others (Lidén & Oxelman, 1996; Dominguez & Wheeler, 1997; Moore, 1998; Benton, 2000; Nixon & Carpenter, 2000).

Supporters of Phylogenetic Nomenclature argue that because Linnaean Taxonomy is based on the concept of rank it is ill-suited to expressing our changing ideas of phylogenetic relationships between species. Rank is problematic because the application of a rigid rank system leads to redundancy and instability. Redundancy is introduced because the Linnaean hierarchy is equidistant: that is to say, every taxon is included in a continuity of ranks from Genus to Kingdom (although this is not stated as mandatory in the Zoological Code). This may be perfectly satisfactory should the phylogeny be perfectly symmetrical (Fig. 3 left). But reconstructed phylogenies are not like this, either because history is genuinely asymmetrical or because of extinctions; they can appear to us as truncated hierarchies (Griffiths, 1973). This means that there are empty ranks (Fig. 3 right). Or to express this in another way: in some parts of the phylogenetic tree the Family rank is equivalent in scope and content to the Order rank elsewhere. Thus, confining oneself to the Recent world, ranks can become redundant in monospecific groups; for example there is nothing more implied by the Family Hominidae than by the Genus *Homo* or the Species *sapiens*.

Rank has been used to imply some level of morphological divergence either in amount or kind. The boundaries of ranks (Genus, Family, Order, etc.) are traditionally, and still usually, judged on morphological divergence. We expect the morphological gaps between Orders to be larger than those between Families, and in turn the latter to be larger than the gaps between Genera. At the same time we expect the variations within Orders than to be greater than those within Families, and these to be greater than the variations within Genera. How much variation and how large a gap is appropriate for families, genera, etc. is usually unstated and is indeed undefinable.

Rank has also been used to signify the kind or quality of divergence to ascribe rank. The action of separating birds as Class Aves equivalent to Class Reptilia is only because of the *kinds* of characters by which birds differ from reptiles. Birds have characters such as wings, feathers and air sacs that enable them to exist in a different adaptive zone. These are deemed by mutual consent to be Class characters and because of this the paraphyletic rubble left behind — the reptiles — also has to have Class status.

So, I have some sympathy with Phylogenetic Nomenclature in the desire to seek a rank-free classification. But there are ways around the problem which do not involve the adoption of a PhyloCode (e.g. Crane & Kenrick, 1997). This is the annotated Linnaean system which by the use of a few conventions (Nelson, 1974; Patterson & Rosen, 1977; Wiley, 1979) can absorb the problems caused by rank yet allow those who wish to retain rank to do so for their own purposes.

It must be remembered that the abolition of ranks can have some rather unfortunate consequences for many people who compile 'diversity indices' based on generic counts or family counts. Here, abolition of rank would immediately affect some palaeontologists and many people studying biodiversity. It is apparently common practice (Dr Sandy Knapp, pers. comm.) in biodiversity inventories to simply note the existence of a representative of a family or genus, because the organism may be new and can only be recognised initially on family characters. With rank abandoned counts are abandoned.

Types/Specifiers

It is difficult to see why Phylogenetic Nomenclature has adopted the new term 'specifier(s)' when, in reality the 'type concept' is still with us, only in a more complicated fashion. The types in Phylogenetic Nomenclature are the specifiers (species, specimens or synapomorphies) *coupled with* a phylogenetic hypothesis. In practice there is no difference between specifiers and types, except that in Phylogenetic Nomenclature it is necessary to cite at least two for every name: node-based (A + B), stem-based (A ← B), apomorphy-based (synapomorphy *a* in A). Under Phylogenetic Nomenclature rules we have the additional complication of the phylogeny, because the name is only to be used within the context of a phylogeny (*PhyloCode, Division 1, Principles, number 6*). As De Queiroz & Gauthier (1992) pointed out, it is always possible to make a mistake about the contents of a clade (the taxa included) and the diagnostics (the characters by which it is recognised) — but it is not possible to make a mistake about the phylogenetic definition. Because of the way names are constructed under phylogenetic taxonomy this must be true. However, it needs to be pointed out that it is not the phylogeny that is important but only the *part* of the phylogeny that is relevant to the name (that portion which includes the specifiers). Other taxa which may have been part of the original phylogeny when the name was erected are free to wander in and out of the named clade. The specifiers and the part of the phylogeny used in erecting the name suffer from the same problem as Linnaean types — they are acting as focal points. How widely or narrowly their naming influence spreads is entirely at the whim of systematists erecting new phylogenies in precisely the same way as in Linnaean taxonomy. Therefore the substituion of types by specifiers (*PhyloCode, Preface*) seems completely unnecessary.

Of course in Linnaean taxonomy the type concept is ultimately tied to characters, attributes of specimens which we can see.

When to name

One of the objections raised against Linnaean Taxonomy is that it is often difficult to name clades without causing a cascade of name changes through rank-ending changes. This, phylogenetic taxonomy claims, will result in clades for which there is much evidence being un-named, and taxonomy becoming out of step with phylogenetic knowledge. Therefore, the unrestricted ability to name clades is seen as an advantage (*PhyloCode, Preface*). However, phylogenetic taxonomy also acknowledges that not all clades need to be named. At first sight this commonsense view may seem odd, considering that the paramount objective of phylogenetic nomenclature is to name clades. Some reasons for naming a clade are given as recommendations in the PhyloCode: '*Criteria that influence the decision whether to name a clade include level of support, phenotypic distinctiveness, economic importance, etc.*' (Preface. paragraph 6). I am not sure what 'etc.' covers, but taking the three that are given I can make some comment.

Level of support. This means that the PhyloCode recommends that we only name clades that are judged to be soundly based with good support. What might this mean? Numbers of synapomorphies, Bremer support, bootstrap support, jackknife support (first order jackknife or second order jackknife), consistency index, retention index, rescaled consistency index, resistance to successive weighting, heavy implied weighting scores, and so on.

Phenotypic distinctiveness. This seems to me to be a curious criterion to use, since much of the PhyloCode's objection against Linnaean Taxonomy is based on the fact that the classical type system does not specify how far from the type the name applies, i.e. how distinctive taxa have to be from a name-bearing type before they become a new genus, family or whatever. Yet, here the PhyloCode seems to be saying the same thing—only in relation to clades. If we are only going to name clades according to phenotypic distinctiveness then this seems to advocate an apomorphy-based definition. We name clades with reference to one or more apomorphies which are judged to be 'significant'. However, apomorphy-based naming is less favoured than the other two definitions because of the subjective assessments of characters; this has been emphasised by Rowe & Gauthier (1992) using the naming of Mammalia as an example.

Economic importance. Well, there are more than enough measures here (e.g. contribution to Gross National Products; financial impact on social conditions, health and welfare; cost-benefit for international aid) but how these are going to be evaluated is difficult.

The point is that the advocates of phylogenetic taxonomy really do not have any more precise reasons for naming a group than do followers of Linnaean Taxonomy and to include advice in the PhyloCode registers a precision which is both unneccesary and undesirable.

How to name

The kind of definition which should be applied in any given clade has been discussed in the context of phylogenetic taxonomy on many occasions (e.g. Lee,

1999b; Sereno, 1999). Here we meet a curiously illogical rationale, since the reason for choosing one kind of definition over another (node-based, stem-based, apomorphy-based) is apparently in order to 'stabilize the taxonomic content of a taxon more than another in the face of local changes in relationships' (Sereno, 1999, p. 329). However, taxonomic content is not the primary purpose of Phylogenetic Nomenclature (*PhyloCode, Division 1. Principles*) and it is therefore unclear why this should be an issue. However, anyone practicing Phylogenetic Nomenclature *must* specify which definition is to be used (see *Practicalities and Governance* above) and therefore some decision has to be made. Several suggestions have been put forward and, for me, the most thorough discussion of this subject is that by Sereno (1999) who advises in which circumstances it may be best to use node-, stem- or apomorphy-based names as well as offering advice on selecting specifiers (types). Despite all the discussion around this subject, the final decision must rest on some estimate as to the resolution, the strength of phylogenetic signal and the potential durability of the phylogeny (crudely put: will those taxa stay in place with the introduction of new data?). In other words some evaluation of the quality of the phylogeny is required. Not surprisingly, Phylogenetic Nomenclature is mute in offering guidelines since there are no agreed criteria amongst the systematic community at large[1]. Therefore while the name of a taxon may well remain stable the applicability of that name within classifications may be decidedly unstable.

There are instances where names can be considered unstable. PhyloCode (Article 15. Conservation) allows that, under certain circumstances involving synonymy and homonomy, authors may apply to the International Committee on Phylogenetic Nomenclature to have names conserved and suppressed. Thus, suppose that with reference to Figure 2A$_2$ Sarcopterygii had been defined as node-based—Sarcopterygii (coelacanth and frog)—and Gnathostomata had been similarly defined (coelacanth and frog). These are clearly synonyms. Let us further imagine that even though date precedence favoured Gnathostomata common usage suggested Sarcopterygii as a more appropriate name (in principle this is similar to the 'prevailing usage' rules of the Zoological Code). As I understand the PhyloCode, Sarcopterygii could be conserved and Gnathostomata suppressed. However, a later author might resurrect Gnathostomata by using a different definition (e.g. Gnathostomata [frog ← lamprey]). This is hardly stability.

Linking a name with a particular phylogeny also leads us into theories of homology, since it is precisely such theories which enable us to recognise the phylogeny in the first place. This is not without difficulty for phylogenetic taxonomy, as may best be explained with reference to the apomorphy-based definition. Historically, in Linnaean Taxonomy apomorphy-based names are those which have caused most confusion, as Rowe & Gauthier (1992) point out in the context of the naming of Mammalia. However, of all of the definitions advocated by phylogenetic

---

[1]There have been many indices devised to try to assess the support of cladograms/phylogenetic trees such as Bremer support, Bootstrap, Jacknife, consistency index, retention index, rescaled consistency index, permutation tail probability tests (and derivatives). All these have their fields of applicability but they are really designed to test the strength of the hierarchial signal, not the stability of the phylogenetic hypothesis which may only be done *a posteriori* with more data. And the problem is compounded if the analysis is carried out under Maximum Likelihood methods — as is often the case with molecular phylogenies — because here the tests applied are undertaken in the context of a particular model of character evolution.

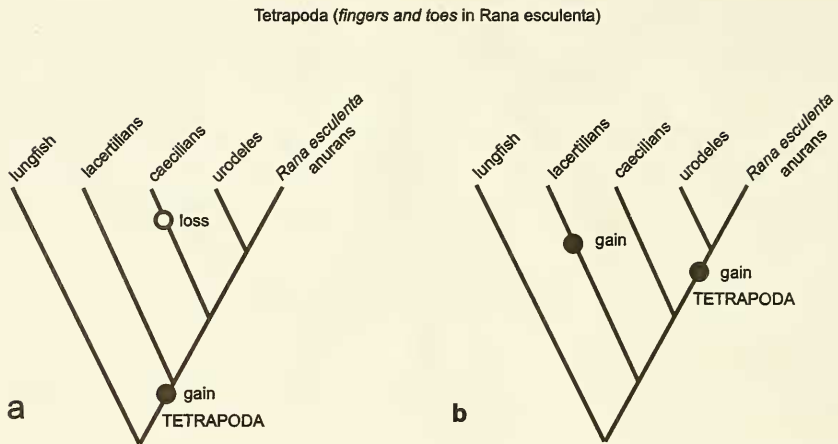Tetrapoda (*fingers and toes* in Rana esculenta)

Figure 4. Apomorphy-based definitions can suffer from alternative equally parsimonious optimisation of potential specifier characters. Here it is suggested that Tetrapods are defined with the specifier fingers and toes. In this phylogeny the taxa lacertilians, urodeles and frogs have fingers and toes but caecilians and lungfish do not. a. Under this optimisation the apomorphy is assumed to have been gained once and the name can be used. b. Alternative optimisation in which fingers and toes are not homologous in lacertilians and the one hand and urodeles+anura on the other.

nomenclature, apomorphy-based naming is the only one which makes specific reference to characters observable in the objects of study (i.e. organisms). But even here there are problems because characters are homologies and homologies are theories. This aspect may not be fully apparent to those taxonomists unfamiliar with phylogenetic systematics. Consider an apomorphy-based definition which may be proposed as 'Tetrapoda is the name given to the clade consisting of all those animals with fingers and toes homologous with those in *Rana esculenta*'. The problem arises over the word homologous. In phylogenetic systematics an homology is a theory and is equivalent to synapomorphy (shared derived character). Let us say that we had arrived at the phylogeny of organisms shown in Figure 4 where lungfishes and caecilians lack fingers and toes whereas lacertilians (mostly), urodeles and *Rana esculenta* have them. There are two ways in which we may imagine the characters 'fingers and toes' to have evolved given this phylogeny. Or, in cladistic terminology, there are two ways of optimising this character on this tree. We could suggest that 'fingers and toes' was gained in the common ancestor of the group lacertilians + *Rana esculenta* and subsequently lost in caecilians. This involves two evolutionary steps (or transformations): one gain and one loss (Fig. 4a). In this case 'fingers and toes' is an homology (shared derived character or synapomorphy) which has been subsequently lost in some members (caecilians) of this group. This type of optimisation is called accelerated transformation (ACCTRAN) because it places the first transformation — no fingers and toes → fingers and toes — at the most inclusive hierarchical level on the tree. One alternative is shown in Figure 4b. Here, it is assumed that 'fingers and toes' is a character that was gained twice — once in lacertilians and again in the common ancestor of the group urodeles + *Rana esculenta*. This optimisation is called delayed transformation (DELTRAN) because it delays the transformation to the most exclusive positions in the hierarchy. In this case 'fingers and toes' is not regarded

as a synapomorphy because it has arisen twice and therefore cannot be considered an homology, and presumably would not be used as an apomorphy-based specifier. However, these two theories of character evolution are equally parsimonious and we would need additional information to choose one alternative as more likely than the other. In order for there to be no ambiguity we need a qualifying phrase to be added to our apomorphy-based definition of Tetrapoda as 'all those animals sharing fingers and toes homologous with those of *Rana esculenta* under the optimising procedure of accelerated transformation'. This is not a particularly utilitarian usage. I admit the example may be contrived: but if fingers and toes were substituted by a particular gene sequence then it may not be easy to argue for or against homology. Perhaps it will be necessary to restate the apomorphy-based definition of Tetrapoda as 'all those animals sharing fingers and toes homologous with those of *Rana esculenta* under any optimising procedure'.

Pain — no gain

The PhyloCode proposes that biologists will gain clarity, efficiency and stability when accepting its premises and adopting its methodology. Advocates of the PhyloCode also claim that these gains are likely to be appreciated by those not interested in phylogeny or nomenclature (Cantino *et al.*, 1999); it is clear that the intention is for the PhyloCode to be understood and used by non-systematists. We need, therefore, to assess what that gain is and at what cost it is to be achieved within the context of biology in general. Cost can only be measured against some standard and therefore some comparison with Linnaean Taxonomy is essential. With respect to clarity and stability there may be no difference between Phylogenetic Nomenclature and Linnaean taxonomy. Within Phylogenetic Nomenclature a name is stable within the context of its specifiers. But so are Linnaean names based on types. The specimen BMNH 1853.11.12.111 is and will remain the name-bearing type (lectotype) of *Clupea harengus*, just as *Struthio camelus* and *Corvus corax* could be regarded as specifiers of the node-based Aves under the PhyloCode. Under Linnaean taxonomy suprafamilial names do not have formal name-bearing types but they may be said to have specifiers. In 1861 T.H. Huxley erected the name Crossopterygii for an assemblage of fossil fishes including *Polypterus, Gyroptychius, Holoptychius, Osteolepis, Dipterus, Phaneropleuron* and *Macropoma*. Two or more of these fishes are the specifiers, if you like, of the name Crossopterygii Huxley, 1861. In the years immediately following Huxley's work the content or membership of the Crossopterygii changed dramatically. This was not surprising because some or all of these fishes were implicated in the ancestry of tetrapods and therefore authors were struggling with a paraphyletic group. However, any scientist foolish enough to struggle with such a group (and I count myself amongst them) is forced to go back to Huxley (1861) to learn the membership of the group and the observations which were used in its recognition. Under phylogenetic taxonomy exactly the same would happen. We cannot gather any relevant details directly from the name Aves (*Struthio camelus* and *Corvus corax*) Joe Doe. We are forced to examine the contents of the clade to understand its membership and presumably we would also be interested in how it was recognised.

When Günther (1871) examined Crossopterygii Huxley, 1861 he decided that the relationships of the included taxa were not as Huxley opined. The phylogeny had

changed and so had the membership of Crossopterygii. But this is precisely what happens under Phylogenetic Nomenclature also (e.g. Sarcopterygii in Fig. 2). If we want to understand the systematic history of a particular taxon we still have to examine all of the phylogenies under which that name has been used because the name itself may be compatible with more than one phylogenetic hypothesis. Thus the claim by phylogenetic taxonomy for clarity and stability within the context of why systematists need the name in the first place is at best illusionary and at worse misleading. There is nothing to be gained.

The pain is administered in several ways. First, for the sake of clarity new names may have to be coined for very familiar groups. The PhyloCode is very clear to point out that this need not be so and suggests that existing names can be redefined under Phylocode conventions by appending a suffix '[P]', meaning that this name is to be used in the sense of phylogenetic taxonomy (Cantino, 2000, p. 87). While this is perfectly feasible, we may ask — will the redefinition be understandable to the many non-systematists who use classifications as their comparative framework? The PhyloCode (Article 11.8) does insist that 'when a clade name is converted from a preexisting genus name or is a new or converted name derived from the stem of a genus name, the definition of the clade name must use the type species of that genus as an internal specifier.'. However, it makes no recommendations as to suprageneric names. Things can go awry. For instance, Laurin (1998) redefined the taxon Anthracosauria under Phylogenetic Nomenclature such that it no longer included its Linnaean type genus *Anthracosaurus*. To use the same name in two completely different contexts will surely lead to confusion, and it puts the onus on the non-systematist to find out the difference or overlap in the meaning of the names. As taxonomists we are hardly serving the wider biological community by this duality and potential confusion.

Second, the PhyloCode is agnostic about characters, relationships, or membership. However, this is precisely the important information which may be of importance to comparative biologists. Thus the retrieval of information may not be as easy as the PhyloCode suggests.

Third, changing hypotheses of relationship will mean that names are used and disused according to the phylogeny in fashion at that time (in Linnaean taxonomy the name will remain the same but the membership may change). This is hardly stability.

Fourth, the PhyloCode names clades, each of which is defined as 'a monophyletic group of species' (PhyloCode, *Preface*). This means that only monophyletic groups be named (there is no other kind of clade). While this is a desirable endpoint we are very far from achieving that phylogenetic resolution. There remain vast branches in the tree of life where monophyly has yet to be demonstrated. Phylogenetic Nomenclature will leave these assemblages of taxa un-named. I find myself in the rather uncomfortable position of being one who agrees strongly that monophyletic groups are the only real biological entities worth consideration and I would never argue for the retention of paraphyletic taxa. But I am also mindful of the fact that for many biologists potentially non-monophyletic groups (e.g. Bryophyta) still serve a useful purpose for their own reasons of communication (say, in ecological studies). Thus we will still have to live with Linnaean names alongside PhyloCode names. The annotated Linnaean system (Wiley, 1979)

can cope with phylogenetic uncertainty to satisfy the systematists without denying names that may be useful elsewhere.

Fifth, adoption of the PhyloCode can and probably would lead to a rapid inflation of names because, quite naturally, individual workers will wish to name the hard-won results of their own phylogenetic investigations. I see this most likely to happen in two areas; molecular systematics and with newly discovered fossil taxa. With respect to the latter de Queiroz & Gauthier (1992, p. 457) recognised this but claimed that since it is palaeontologists who are most concerned with phylogenies they should live with this problem, which they dismissed as minor since 'there are already more taxon names than anyone can remember — then naming clades seems preferable to leaving them unnamed . . .'. Thus, in one sense, phylogenetic systematists get what they deserve. But in another sense, phylogenetic systematists are not serving the wider biological community by introducing a plethora of names, each with their own definitions which need to be understood before they can be used by others.

Conclusion

The intention of the PhyloCode is to name clades and it is therefore free of empirical content (with the possible exception of the apomorphy-based definition). In trying to name hypotheses the PhyloCode puts the onus on the users of the names to assess the confidence we may have in one particular clade or another before selecting a name that matches that choice. Users of Linnaean taxonomy are, of course, forced to do the same, but no name changes need be required. The alleged clarity, efficiency and stability claimed by the PhyloCode do not stand critical examination and it needs to be asked what exactly has been gained. More importantly the biological community will have to judge whether the alleged gains are worth the undoubted pain.

Acknowledgements

References

Benton, M.J. 2000. Stems, nodes, crown clades, and rank-free lists: is Linnaeus dead? *Biological Reviews*, **75**: 633–648.

Cantino, P.D. 2000. Phylogenetic nomenclature: addressing some concerns. *Taxon*, **49**: 85–93.

Cantino, P.D., Byrant, H.N., De Queiroz, K., Donoghue, M.J., Eriksson, T., Hillis, D.M. & Lee, M.S.Y. 1999. Species names in phylogenetic nomenclature. *Systematic Biology*, **48**: 790–807.

Crane, P.R. & Kenrick, P. 1997. Problems in cladistic classification: higher-level relationships in land plants. *Aliso* **15**: 87–104.

De Queiroz, K. & Gauthier, J. 1990. Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Systematic Zoology*, **39**: 307–322.

De Queiroz, K. & Gauthier, J. 1992. Phylogenetic taxonomy. *Annual Review of Ecology and Systematics*, **23**: 449–480.

De Queiroz, K. & Gauthier, J. 1994. Towards a phylogenetic system of biological nomenclature. *Trends in Ecology and Evolution*, **9**: 27–31.

Dominguez, E. & Wheeler, Q.D. 1997. Taxonomic stability is ignorance. *Cladistics,* **13**: 367–372.

Forey, P.L. 1980. *Latimeria*: a paradoxical fish. *Proceedings of the Royal Society of London,* B, **208**: 369–384.

Griffiths, G.C.D. 1973. Some fundamental problems in biological classification. *Systematic Zoology,* **22**: 338–343.

Günther, A.C.L.G. 1871. Description of *Ceratodus,* a genus of ganoid fishes, recently discovered in the rivers of Queensland, Australia. *Philosophical Transactions of the Royal Society of London,* **161**: 511–571

Huxley, T.H. 1861. Preliminary essay upon the systematic arrangement of the fishes of the Devonian epoch. *Memoirs of the Geological Survey of the United Kingdom,* **Decade 10**: 1–40.

Jefferies, R.P.S. 1979. The origin of chordates — a methodological essay. Pp. 443–477 *in* House, M.R. (Ed.), The origin of major invertebrate groups. *Systematics Association Special Volume Series,* no. 12.

Lagios, M.D. 1979. The coelacanth and the Chondrichthyes as sister groups: a review of shared apomorph characters and a cladistic analysis and reinterpretation. *Occasional papers of the California Academy of Sciences,* **134**: 25–44.

Laurin, M. 1998. The importance of global parsimony and historical bias in understanding tetrapod evolution. Part 1. Systematics, middle ear evolution and jaw suspension. *Annales des Sciences Naturelles, Zoologie et Biologie Animale* **19**: 1–42.

Lee, M.S. 1998. Phylogenetic uncertainty, molecular sequences, and the definition of taxon names. *Systematic Biology,* **47**: 719–726.

Lee, M.S.Y. 1999a. Reference taxa and phylogenetic nomenclature. *Taxon,* **48**: 31–34.

Lee, M.S.Y. 1999b. Stability of higher taxa in phylogenetic nomenclature — some comments on Moore (1998). *Zoologica Scripta,* **28**: 361–366.

Lidén, M. & Oxelman, B. 1996. Do we need 'phylogenetic taxonomy'? *Zoologica Scripta,* **25**: 183–185.

Moore, G. 1998. A comparison of traditional and phylogenetic nomenclature. *Taxon,* **47**: 561–579.

Nelson, G. 1974. Classification as an expression of phyogenetic relationships. *Systematic Zoology,* **22**: 344–349.

Nixon, K.C. & Carpenter, J.M. 2000. On the other 'Phylogenetic Systematics'. *Cladistics,* **16**: 298–318.

Patterson, C. & Rosen, D.E. 1977. Review of ichthyodectiform and other Mesozoic teleost fishes and the theory and practice of classifying fossils. *Bulletin of the American Museum of Natural History,* **158**: 81–172.

Rowe, T. & Gauthier, J. 1992. Ancestry, paleontology, and definition of the name Mammalia. *Systematic Biology,* **41**: 372–378.

Schultze, H.–P. 1987. Dipnoans as sarcopterygians. Pp. 39–74 *in* Bemis, W.E., Burggren, W.W. & Kemp, N. (Eds.), *The biology and evolution of lungfishes.* Alan R. Liss, New York.

Sereno, P.C. 1999. Definitions in phylogenetic taxonomy: critique and rationale. *Systematic Biology,* **48**: 329–351.

Wiley, E.O. 1979. An annotated Linnaean hierarchy, with comments on natural taxa and competing systems. *Systematic Zoology,* **28**: 308–337.